

## Coverpage

Title: Defending the traditional interpretations of Kant's Formula of a Law of Nature

Abstract: In this paper I defend the traditional interpretations of Kant's Formula of a Law of Nature from recent attacks leveled by Faviola Rivera-Castro, James Furner, Ido Geiger, Pauline Kleingeld and Sven Nyholm. After a short introduction, the paper is divided into four main sections. In the first, I set out the basics of the three traditional interpretations, the Logical Contradiction Interpretation, the Practical Contradiction Interpretation and the Teleological Contradiction Interpretation. In the second, I examine the work of Geiger, Kleingeld and Nyholm: these three commentators reject the traditional interpretations entirely, but I argue that this rejection is ill-founded. In the third and fourth, I take a detailed look at Furner's work, work in which he seeks to revise (rather than reject) the traditional interpretations. I argue that, despite his more modest aims, Furner's revision is also ill-founded.

Keywords: Categorical Imperative; Contradiction in Conception; Contradiction in Willing; Formula of a Law of Nature; Kant's Ethics; Kantian Ethics; Logical Contradiction Interpretation; Practical Contradiction Interpretation

Note: published  
version underwent  
revisions

## **Defending the traditional interpretations of Kant's Formula of a Law of Nature**

In two recent articles, James Furner has offered a novel and interesting interpretation of Kant's Formula of a Law of Nature, an interpretation that he says is truer to the letter and the spirit of Kant than the traditional interpretations it revises. In this paper I argue that although Furner's interpretation fails on its own terms, seeing why it does so is illustrative of important aspects of Kant's work and recent Kant scholarship. In particular, I intend to show that the traditional interpretations of Kant's Formula of a Law of Nature are more powerful than often supposed.

The paper is divided into four sections. In the first, I set out some of the background required to understand Furner's project: the traditional interpretations mentioned in the previous paragraph. In the second, I defend these interpretations against attacks from Ido Geiger, Pauline Kleingeld and Sven Nyholm, all of whom seek to replace them wholesale. In the third, I concentrate on Furner's (revisionary) interpretation of the so-called Contradiction in Conception test. In the fourth, I concentrate on Furner's (revisionary) interpretation of the so-called Contradiction in the Will test. Along the way, I also confront Faviola Rivera-Castro, who has offered competing revisionary interpretations of these tests.

### **Section 1. The traditional interpretations**

The task of the *Groundwork for a metaphysics of morals* is to articulate the Supreme Law of Morality.<sup>1</sup> In particular, Kant wants to articulate this law as it appears to humans. According to Kant, humans are imperfectly rational agents: we are subject to various limitations and hin-

drances that tempt us away from doing our duty.<sup>ii</sup> Because of this, the Supreme Law of Morality manifests to us as an imperative: the Categorical Imperative (CI).<sup>iii</sup>

Kant formulates the CI in various ways. The formulation that is most important for current purposes is called the Formula of a Law of Nature (FLN). Kant introduces this formulation in section II of the *Groundwork for a metaphysics of morals*. It runs as follows: '*act thusly, as if the maxim of your action through your will should become a **universal law of nature***.'<sup>iv</sup> Kant argues that there are two ways that a maxim can fail the test of FLN:

Some actions are so constituted, that their maxim cannot even [einmal] be *thought* without contradiction as a law of nature; let alone [weit gefehlt], that one yet could *will*, it *should* become such a one. With others this inner impossibility is indeed not to be found, but it is nevertheless impossible, to *will*, that their maxim be raised to the universality of a law of nature, because such a will would contradict itself.<sup>v</sup>

Kant then tells us that these two ways of failing FLN can be used to carve up the moral landscape into more traditional ethical categories, gesturing toward a task that he reserves for his future *Metaphysics of morals*.<sup>vi</sup>

In her seminal *Acting on Principle*, Onora O'Neill introduces some now widely-used terminology to refer to these two ways of failing FLN.<sup>vii</sup> She calls the first kind of contradiction, the 'inner impossibility' of a maxim that cannot be thought as a law of nature, a Contradiction in Conception (CC) (2013: 143ff). She calls the second kind of contradiction, when a will that rais-

es a maxim to the universality of a law of nature contradicts itself, a Contradiction in Willing (CW) (2013: 173ff). She also introduces the notion of a maxim's Universalized Typified Counterpart (UTC), shorthand for a maxim raised to the universality of a law of nature (2013: 141ff).

Christine Korsgaard then organizes the interpretive debate about these tests in her similarly influential 'Kant's Formula of Universal Law.'<sup>viii</sup> Korsgaard distinguishes between three different interpretations of the CC test and the CW test: the logical contradiction interpretation (LCI), the practical contradiction interpretation (PCI) and the teleological contradiction interpretation (TCI).

The differences between these three interpretations can be explained by means of the CC test. On the LCI, a maxim cannot be thought as a UTC because the action in the maxim would be impossible if the maxim were a UTC.<sup>ix</sup> For example, consider the maxim 'to tell a lying promise in order to get some ready money when in financial distress.'<sup>x</sup> In a world in which this maxim is a UTC, lying promises are impossible because the convention of promise-keeping would not be possible. Thus, this maxim fails the LCI.<sup>xi</sup>

On the PCI, a maxim cannot be thought as a UTC because it would not be possible to attain the agent's end by means of the action if it were so.<sup>xii</sup> Thus, the maxim discussed in the previous paragraph also would fail the PCI, but it would do so for a different reason: it would fail the PCI because an agent would not be able to attain ready money by means of a lying promise if that maxim were a UTC.<sup>xiii</sup> That said, the reason the action is ineffectual is that the convention of promise-keeping would not be possible, whence it may be seen that a maxim that fails the LCI (*a fortiori*) will fail the PCI.<sup>xiv</sup>

On the TCI, the question to be asked is whether the UTC of a maxim can be willed as a law of a teleologically organized nature.<sup>xv</sup> Thus, the lying promise maxim would fail the TCI. But this time the reason would be that its UTC could not function as a law of a teleologically organized nature.

These remarks now can be extended to the CW test. For example, one way of explaining the LCI of the CW test would be to say that the UTC requires the existence (or nonexistence) of some *X* whereas some essential end of the agent requires the nonexistence (or existence) of *X*. Similarly, one way of explaining the PCI of the CW test would be to say that the UTC would render it impossible for some essential purpose of the agent's to be fulfilled. And one way of explaining the TCI of the CW test would be to appeal to the harmony of ends in a teleologically organized nature and to show that the UTC would undermine this harmony. In all three of these cases, the impossibility is not an 'inner' one, a problem between the UTC and the original maxim: the impossibility arises from conjoining the UTC with consideration of some other essential end of the agent (LCI or PCI) or with some sort of overarching harmony of ends (TCI). This will be explained more in section 4.

## **Section 2. Kleingeld's, Nyholm's and Geiger's rejections of the traditional interpretations**

As noted above, Furner endorses a revised version of the LCI. But some interpreters, like Kleingeld, Nyholm and Geiger, take a more radical line of attack: they seek to show that these interpretations are fundamentally misguided. Thus, I would like to say a few words about these more radical lines of attack before turning to Furner.

Kleingeld seeks to replace the traditional interpretations with her own, the Volitional Self-Contradiction Account (VSCA). To do so, she notes that FLN is supposed to be equivalent to another version of the CI, the Formula of Universal Law (FUL): '*act only according to that maxim, through which you simultaneously can will, that it become a universal law.*'<sup>xvi</sup> Kleingeld emphasizes the simultaneity clause of FUL and the fact that it is absent from the traditional interpretations:

On the Volitional Self-Contradiction account, the fact that one's purpose is thwarted or that one's project is impossible when the maxim is a universal law does not constitute the relevant contradiction but rather produces it. We cannot will the maxim and simultaneously will that it be a universal law because the maxim cannot be a universal law. (2017: 107)

Kleingeld's idea here is that the PCI (purpose being thwarted) and the LCI (project being impossible) reveal that a maxim cannot be a universal law, and this in turn shows that we cannot will the maxim and simultaneously will that it be a universal law (VSCA). I would like to say two things about this.

First, there are suppressed premises in Kleingeld's reasoning. Here are two: (1) we cannot will a maxim as a universal law if it cannot be a universal law, and (2) we cannot will a maxim and simultaneously will that it be a universal law if we cannot will it as a universal law. (1) is nontrivial but gets into issues too large to grapple with here.<sup>xvii</sup> The more immediate problem is

that (2) seems to render the simultaneity condition superfluous, exactly the opposite of what Kleingeld wants.<sup>xviii</sup>

Second, the LCI and the PCI generate contradictions in conception by showing that an agent cannot will a maxim in a world governed by the maxim's UTC. Thus, these interpretations work by showing that we cannot will a maxim and simultaneously will its UTC. So I do not think the VSCA is materially different from the traditional interpretations (*pace* Kleingeld).<sup>xix</sup>

Indeed, Kleingeld's explanation of the VSCA seems to echo O'Neill's summary of the PCI. Consider the following passage from Kleingeld:

...the simultaneity condition, spelled out in full, is a condition of being simultaneously willable: the possibility of willing a maxim as one's own action principle (as willed independently of or prior to its universalization) and simultaneously willing this maxim as a universal law, without this generating a volitional self-contradiction. (2017: 97)

Now compare this with O'Neill: 'To universalize maxims agents must satisfy themselves that they can both adopt the maxim and simultaneously will that others do so' (1990: 95).<sup>xx</sup> The two are strikingly similar, rendering Kleingeld's goal of showing the superiority of her interpretation at least *prima facie* questionable. If I am right, then Kleingeld's VSCA does not enjoy any substantial advantages over the traditional interpretations, a thesis I shall return to in section 4 below.<sup>xxi</sup>

Nyholm's attempt to replace the traditional interpretations, which faces a different problem, is built up in three steps. First, he argues that the word 'nature' in FLN refers to a thing's constitution. On these grounds he asserts that FLN is instructing agents to act only on maxims that can function as laws governing their distinctive constitutions (2015: 292).<sup>xxii</sup> Second, he argues that as imperfectly rational beings, there are two distinctive features of our constitutions: (a) the ability to govern ourselves in accordance with a law, and (b) our aim for our own happiness (2015: 293). Therefore, third, for a maxim to pass FLN it must be possible for us fully to realize these two features in a world in which it is a law of nature. That is, it must be possible for us to be fully autonomous and perfectly happy in such a world. Nyholm summarizes this as follows:

...as Kant sees things, basic principles fit to serve as universal laws...are ones under which we could all potentially both be happy and, at the same time, be preserved and function as beings autonomously governed by their own reason. More simply put, the basic principles fit to be universal laws are ones under which—if and insofar as the Universe functions in a way favorable to the ends of morality—happiness and virtuous autonomy would both be possible for us all... (2015: 294)

The main problem with this account, however, is that according to Kant, moral laws just are the laws of autonomy, and so the happiness constraint cannot play the role Nyholm assigns to it.<sup>xxiii</sup> To put this another way, on Kant's account a world in which agents are following the laws of autonomy just is a world in which agents are following the Supreme Law of Morality.

This problem might be overcome if the connection between morality and happiness were analytic. But such a rejoinder is not open to Nyholm: as a reading of Kant, this would run aground on Kant's explicit (and repeated) contention that the connection between morality and happiness, if there is one, must be synthetic.<sup>xxiv</sup> Thus, Nyholm's interpretation cannot be correct.

I shall wrap up this section with a brief examination of Geiger. Geiger objects to the idea that 'substantive moral permissions and proscriptions can be generated from a formal universalization test without any moral knowledge as input,' an idea that seems to be at the root of the traditional interpretations (2010: 272). According to Geiger, the universalizability tests cannot be used to generate knowledge: 'the agents Kant has in mind do not ordinarily learn anything they did not know when they apply the FUL' (2010: 281). The primary aim in articulating these tests is, on Geiger's reading, to 'state and substantiate the formal formulation of Kant's conception of moral motivation and action' (2010: 272).

However, I think Geiger's account faces both exegetical and philosophical difficulties. Exegetically, Geiger has a hard time accounting for Kant's repeated claims that the universalizability tests should be used as a compass to navigate the moral landscape.<sup>xxv</sup> Moreover, philosophically it is hard to see how the CI could be limited in the way in which Geiger suggests.

It is true that in the *Groundwork for a metaphysics of morals* Kant does not seem to take himself to be saying anything surprising in illustrating how maxims of suicide, false-promising, idleness and non-beneficence fail FLN. But that does not impugn the role of FLN as a test of maxims. Quite the contrary: in trying to show (with these examples) that FLN generates results that conform to pre-theoretic intuition Kant is illustrating the basic meaning of FLN by means of examples. Stating a formal constraint on moral motivation and action is consistent with and, in-

deed, undergirds the project of articulating a test wherewith to generate new moral knowledge. As sections 1 and 2 of this paper have shown, Kant's readers have extrapolated this basic meaning in different ways. But to assert that such interpretive projects are confused is, in my view, mistaken: if agents ought to act only on maxims that they can will as laws of nature, then any maxim that cannot be so willed is impermissible. Kant gives us hints as to how to determine whether this is so, and those hints are what the LCI, PCI and TCI are extrapolated from. So although I now shall argue that Furner's attempt to revise the LCI fails, I agree with his basic outlook (*pace* Geiger).

### **Section 3. Furner's interpretation of the CC test**

Furner offers what he calls a Causal-Teleological version of the LCI (CTLCI). The two essential features are as follows. First, the CTLCI supplements the traditional LCI by allowing for the contradiction to manifest in the impossibility of an agent attaining the end proposed in a maxim in a world in which the maxim is a UTC (2017a: 10). This is in addition to (not instead of) the more traditional LCI impossibility of the agent performing the action in a world in which the maxim is a UTC. Thus, the CTLCI joins the LCI with the PCI.<sup>xxvi</sup>

Second, the CTLCI is distinguished from other versions of the LCI in that it appeals to a teleological principle. In particular, it appeals to the idea that organized beings (like humans) are perfectly endowed for the pursuit of their ends (2017a: 5). Moreover, it appeals to the idea that agents' maxims are to be regarded as an indispensable part of their moral lives. Thus, in addition to joining the LCI with the PCI, the CTLCI joins the LCI with the TCI.

Furner summarizes these two features of the CTLCI in the following passage:

On this view, to imagine a maxim as a universal law of nature is to imagine a hypothetical world in which (1) everyone who satisfies its condition cannot but adopt its means and achieve its end ('causal'); and in which everyone regards the maxim as an indispensable part of their moral life, and regards their endowments as suiting them to lead a moral life that includes it ('teleological'). (2017a: 3)

What I have described as the first essential feature of Furner's CTLCI, in which he unites the LCI with the PCI, is what he describes in this passage as (1): not only must the action (the 'means' in the maxim) be performable if the maxim's UTC is a law of nature, but also the end at which the agent is aiming must be attainable by that action if the maxim's UTC is a law of nature.<sup>xxvii</sup> What I have described as the second essential feature of Furner's CTLCI, in which he unites the LCI with the TCI, is what he describes in this passage as (2): the maxim to be tested is to be regarded as an essential part of the moral life in a world in which endowments are perfectly suited for that moral life.

Furner also elaborates on a nonessential feature of his CTLCI. This feature is nonessential in the sense that although Furner believes that it is true, he says that someone could subscribe to the CTLCI even if s/he disagreed. It is that 'a maxim only fails the CC test if a logical contradiction is produced in *every* hypothetical system of nature to which its counterpart universal law of nature could consistently belong' (2017a: 8). That is, it will not suffice, according to Furner, if there is *some* hypothetical system of nature conceived using the maxim's UTC that manifests a

contradiction; for a maxim to fail the CC test (and, thus, be impermissible), *every* such system must be contradictory.<sup>xxviii</sup>

Now Korsgaard maintains that there is textual evidence in favor of the LCI, PCI and TCI. Because of this, she says she is going to argue for the PCI on *philosophical* grounds. In fact Korsgaard argues that any interpretation must be based, at least in part, on philosophical grounds:

...no interpretation can be based on textual considerations alone. Language supporting all of them can be found in Kant's texts, and it seems possible that he was not aware of the differences among them. (1996: 80)

Accordingly, Korsgaard ends up arguing that the PCI handles some maxims involving violent natural actions like killing (as opposed to conventional actions like promising) better than the LCI (1996: 97-101). Similarly, she argues that the TCI seems to presuppose a purposiveness that the agent willing a particular maxim is ready to forswear, thus rendering dubious any attempt to derive a contradiction or impossibility in the agent's will (1996: 90).

One might question Korsgaard's approach. For one thing, the mere fact that language supporting all three interpretations can be found does not indicate that the textual evidence supporting all three interpretations is of equal weight. Perhaps a careful consideration of both direct textual support and indirect doctrinal support will tip the scales in one direction or the other. For another thing, if textual evidence really does underdetermine the interpretation of the CC test, one might think that we should remain agnostic with regard to which interpretation is Kant's.

The principle of charity does not seem to require us to ascribe to Kant the interpretation that we find most philosophically attractive when the texts are ambiguous.

But whatever the (de)merits of Korsgaard's approach, it is not the approach favored by Furner.<sup>xxxix</sup> Furner argues for the CTLCI solely on exegetical grounds.<sup>xxx</sup> In particular, he offers three reasons for preferring his reading over the LCI, PCI and TCI:

1. The CTLCI handles Kant's suicide case better than the TCI, PCI and traditional versions of the LCI (2017a: 11-16).
2. The CTLCI handles Kant's lying promise case better than the PCI (2017a: 16-18).
3. The CTLCI gives a unified account of the suicide case and lying promise case (2017a: 2 and 19).

Let us examine each of these, briefly, in turn.

Kant's suicide case is the first one he attempts to handle with FLN.<sup>xxxix</sup> Kant imagines an agent acting on this maxim: 'I make it to myself from self-love into a principle, when life, with its long protraction, threatens more evil than it promises pleasantness, to cut it short for myself.'<sup>xxxiii</sup> Kant argues that such a maxim fails the test:

...one soon sees however, that a nature, whose law it was, through this same sensation, whose determination it is, to drive toward the promotion of life, to destroy life itself, would contradict itself and thus would not subsist as nature...<sup>xxxiii</sup>

On Furner's rendition of this case, a contradiction arises because the suicide maxim is to be conceived as an indispensable part of a hypothetical agent's moral life, whence it follows that such an agent 'must adopt all necessary means available to them to preserve their capacity to act on it' (2017a: 13). These necessary means include the agent's 'capacity for 'sensation', which records pleasures and pains, and so allows them to judge when it is necessary to end their life to make it bearable until the end' (2017a: 13). Furner concludes that the suicide maxim fails the CC test because it is impossible for an agent both to end his/her life and to preserve his/her capacity for sensation.<sup>xxxiv</sup>

However, the contradiction alleged on the CTCLI is illusory.<sup>xxxv</sup> The necessary means to acting on the suicide maxim do not include the unconditional preservation of the capacity for pleasure and pain. If the maxim is 'to reduce suffering by cutting life short,' then the necessary (and sufficient) means are a peaceful, relatively painless death. There are plenty of straightforward ways to achieve that, and there is no CTCLI contradiction that will arise from their adoption in a world in which the UTC of this maxim holds, at least not on Furner's argument. So Furner's argument for 1, I think, fails.<sup>xxxvi</sup> I turn now to 2.

I have discussed the lying promise case already in section 1 when setting out the differences between the LCI, PCI and TCI. So rather than rehash that, I shall move straight into the CTCLI's version of the lying promise case and, then, Furner's objections to the PCI's version of this case.

According to Furner, in a world in which the UTC of the lying promise maxim holds, 'everyone believes...that those in need of money may properly have no intention to keep to what they say about repaying it' (2017a: 17). Because of this, Furner says that it would be 'logically

contradictory to imagine a world' in which everyone in need of money makes a false promise and obtains a loan, for in such a world 'no one in need of money is understood as requesting a loan and so no one who inquires receives a loan' (2017a: 17).

Furner cites two reasons why his interpretation of this case is better than that of the PCI. First, he points out that according to Korsgaard, if the UTC of the lying promise maxim held, then the 'practice of offering and accepting promises would have died out' (Korsgaard 1996: 82; quoted on Furner 2017a: 17). Furner says that this is problematic because Kant says that the lying promise maxim, 'so soon as it were made into a universal law, would have to destroy itself.'<sup>xxxvii</sup> That is, the time delay required by Korsgaard conflicts with the text. Second, Furner points out that in the text, Kant says that the lying promise maxim would make 'the promise and the end...impossible.'<sup>xxxviii</sup> The PCI, by way of contrast, focuses only on the end rather than the promise itself. The PCI thus is unable to account for the fact that 'Kant mentions both, on a par with one another' (Furner 2017a: 18).

However, neither of Furner's reasons withstands critical scrutiny. The main problem with Furner's first point is that the part of Korsgaard's reading at which Furner is taking aim is inessential to the PCI.<sup>xxxix</sup> This becomes evident once we understand the rationale behind it: the rationale is that it does not suffice for the UTC of the lying promise maxim to hold for the convention of promising to disappear; in addition, everybody must know that this UTC holds. A proponent of the PCI could adopt a publicity condition instead of going Korsgaard's route, but this is something with which the CTLCI must grapple too, so there is no advantage here to either side.

The problem with Furner's second point is simpler. The PCI has a ready explanation of the fact that Kant mentions that both the promise and its purpose would be impossible: he does so because the impossibility of the promise is what renders the purpose impossible, and the latter is what matters. It is in fact versions of the LCI that seem to face a problem here, for they must explain Kant's mention of the impossibility of the purpose, something that is irrelevant to them. Korsgaard takes this to be too thin a reed on which to build the case for the PCI. But the point for present purposes is that Furner's arguments for 2 also fail. So I turn, finally, to 3.

Furner's argument regarding 3 is as follows: (A) the CTLCI gives a unified account of the suicide case and the lying promise case; (B) none of the other interpretations does so; therefore, (C) the unity offered by the CTLCI is an advantage. The problem, however, is that each of these assertions is questionable or false.

There are two issues with premise (A): first, as discussed above, the CTLCI is not able to handle the suicide case (*pace* Furner); and second, even if the CTLCI account of the suicide case worked, it may be seen that it is relevantly different from the CTLCI account of the lying promise (whereas the first relies on teleological assumptions about nature, the second relies on the fact that promising depends on a convention). Premise (B) is no better off: Furner fails to notice that Paton's TCI *is* able to give a unified account of these two cases. Finally, there is a questionable tacit premise employed in the move from (A) and (B) to (C): that all else being equal, an interpretation of the CC test *should* give a unified account of these cases. Given the philosophical differences manifest in *Kant's* accounts of these two cases, this tacit premise is suspect.<sup>xI</sup> So I shall put this down now to focus on Furner's account of the CW test.<sup>xII</sup>

#### **Section 4. Furner's interpretation of the CW test**

Furner calls his interpretation of the CW test the Extravagant Imperfect Nature Interpretation (EINI). Like the CTLCI, the EINI has two defining features. The first is that it 'interprets the CW test as a test of a contradiction in the will of a hypothetical creator' (2017b: 315). This distinguishes the EINI from the PCI, LCI and TCI, all of which take the contradiction to be in the will of the agent proposing to act on the maxim being tested. The second is that it hinges on the idea that 'a hypothetical creator wills a non-extravagant perfect nature' (2017b: 315).

According to Furner, a nature is extravagant if its agents have a capacity that is 'redundant in respect of their ends' (2017b: 316). He illustrates this by appeal to the way the human appendix has been viewed. Thus, hypothetical agents have extravagant natures if that nature 'includes a capacity that need not be exercised or even developed for them to pursue the ends they are to pursue' (2017b: 317). Similarly, according to Furner a hypothetical agent would be imperfect if its nature 'includes a capacity sub-optimal for these ends' (2017b: 309).

As with the CTLCI, Furner says that his project is 'strictly interpretative' (2017b: 309), and he offers three reasons for preferring EINI over other interpretations of the CW test:

1. EINI satisfies assessment criteria which other interpretations fail (2017b: 309-318).
2. EINI offers a more plausible reading of Kant's natural talents example than other interpretations (2017b: 319-321).
3. EINI offers a more plausible reading of Kant's beneficence example than other interpretations (2017b: 321-322).

And as with Furner's reasons for CTLCI, I shall examine each of these, briefly, in turn.

Furner formulates four assessment criteria which any interpretation of the CW test should meet. According to Furner, the CW test confirms a maxim to be permissible on account of (i) two contradictory acts of willing (ii) belonging to the same will (iii) that wills a maxim become a universal law of nature (iv) while eschewing external premises about empirical ends (2017b: 311). For the sake of brevity I shall confine myself to Furner's claim that the PCI fails to satisfy (ii), (iii) and (iv) whereas the EINI satisfies them.

Regarding (ii) and (iii), Furner notes that 'FLN asks you to imagine that your maxim *'become by your will a UNIVERSAL LAW OF NATURE'*' (2017b: 312). Furner points out that no finite rational being can believe they have the power to make their maxim into a universal law of nature and, thus, that it makes no sense to ascribe the UTC of a maxim as an act of will to a finite rational being. On these grounds, Furner maintains that the PCI does not satisfy (ii) or (iii) but that the EINI does: only a hypothetical creator can will the UTC of a maxim, so only the EINI is able to locate the two contradictory acts of willing in a single will (2017b: 312-313 and 315-317).

There are two problems with Furner's argument here. The first problem, which I shall return to below when looking at Furner's discussion of the natural talents example and the beneficence example, is that although a hypothetical creator might be able to will the UTC of a maxim, such a creator's ability to will anything that would contradict this UTC is quite attenuated. For instance, unlike a finite rational being, a hypothetical creator, *qua* creator, need not will to develop any of his/her talents (they are presumably already perfect) even though, *qua* creator, s/

he would be able to will a UTC like 'nobody ever will develop any of their talents.' Similarly, unlike a finite rational being, a hypothetical creator, *qua* creator, need not ever will to receive help (s/he presumably never will need it) although, *qua* creator, s/he would be able to will a UTC like 'nobody ever will help anybody.' This, which is the flip-side of Furner's point about willing a UTC, makes it difficult for EINI to locate two contradictory acts of willing in the same will.

The second problem is that Furner has misrepresented Kant's text. It is perhaps true that no (non-delusional) finite rational agent genuinely can will the UTC of a maxim. But any finite rational agent presumably can act '*as if*' his/her maxim is a universal law of nature, and the point of FLN is that this can reveal contradictions in the agent's will, contradictions that render that agent's will bad. That is why, in the section of FLN which Furner quotes to make his point, the maxim is to become a universal law of nature through '*your*' will: it is the will of the agent at which FLN is directed, not the will of a hypothetical creator. I turn now to (iv).

Furner gives three grounds for the idea that any interpretation of the CW test should eschew 'external premises about empirical ends.' First, allowing such ends into the argument would compromise the fact that the CI 'necessitates the will immediately and unconditionally' (2017b: 311). Second, any such ends would be contingent and, thus, it would be possible to conceive a rational species without it: such a species therefore would not have the same duties as we (2017b: 311). And third, the ground for following FLN is *a priori* and necessary, and this would be undermined by allowing empirical ends into the mix (2017b: 312).

However, there are at least two problems with all of this. The first is that Furner mistakes the way in which the PCI appeals to empirical ends. As noted at the end of section 1 of this paper, as the PCI interprets the CW, the UTC of a maxim would render it impossible for some es-

sential purpose of the agent's to be fulfilled, whence it follows that the maxim cannot be willed as a universal law of nature (thus evincing a contradiction in the agent's will when s/he tries to act, following FLN, as if his/her maxim were such a law). This does not compromise the bindingness of the CI, FLN or the CW test, nor does it impugn the *a priori* grounds which, according to Kant, we have for conforming to them. Rather, it tells us how to apply them to the details of our empirically messy, actualized lives.<sup>xlii</sup>

This *does* mean that, according to the PCI, what *I* can will as a universal law of nature and, thus, what FLN commands *me* to do might differ from what a member of some *other* rational species can will as a universal law of nature and, thus, what FLN commands *him/her/it* to do. But that leads me to the second problem, which is that there is textual support for this sort of thing in Kant.<sup>xliii</sup>

Here are two pieces of direct textual support in favor of this: (I) according to Kant, perfectly rational beings do not have any duties: the CI is directed specifically to imperfectly rational beings, and God, for example, is 'a being, which has nothing but right and no duty';<sup>xliv</sup> and (II) in Kant's derivation of duties in the *Metaphysics of morals*, he appeals to various 'external premises' with empirical information, including information about ends. For example, Kant's derivation of the duty of beneficence is based on the fact that 'I will the goodwill (*benevolentiam*) of everyone else toward [gegen] me.'<sup>xlvi</sup>

Here is a piece of indirect, doctrinal support in favor of this: Kant is committed to the principle 'ought implies can' (OIC). Now consider the following binary:

- If I had the power to end world suffering by snapping my fingers, I ought to do so;

- I do not have a duty to end world suffering by snapping my fingers. (Graham 2011: 372)

What I take this binary to illustrate is that commitment to OIC brings along with it commitment to the idea that duties are not invariant across rational beings with vastly different abilities: some external premises with empirical information are needed in order to determine the results of FLN as applied to an individual. The following passage may be adduced to show that Kant would agree with this: 'it is not enjoined on man [that] he should...foreswear his natural end, happiness; for this he cannot [do].'<sup>xlvi</sup>

Based on this evidence, I think that Furner's argument about (iv) is misguided.<sup>xlvii</sup> I conclude that the PCI is able to weather Furner's assessment criteria: Furner's reason 1 for supporting the EINI does not connect. I turn now to reasons 2 and 3.

Furner's reasons 2 and 3 concern Kant's natural talents example and his beneficence example, respectively. These are the third and fourth examples Kant uses to illustrate FLN (after the suicide example and the lying promise example). The natural talents example involves an agent who is in comfortable circumstances. The agent is contemplating a life of idleness and enjoyment instead of cultivating a talent which 'could make him into a useful person for [in] all sorts of purpose.'<sup>xlviii</sup> Kant maintains that such a maxim would fail the CW test because 'as a rational being he necessarily wills, that all capacities in him be developed, because they are serviceable and given to him nevertheless for [zu] all sorts of possible purposes.'<sup>xliv</sup>

The beneficence example involves another agent in comfortable circumstances. But this one is contemplating a maxim never to help anyone despite the fact that 'he sees, that others have to struggle with great hardships.'<sup>l</sup> Kant argues that such a maxim would fail the CW test because

even for an agent in comfortable circumstances a case very often can arise in which 'he needs the love and participation of others...[and yet] would rob himself of all hope of support, which he wishes for himself' by virtue of willing his maxim of indifference as a universal law of nature.<sup>li</sup>

I would like to say three things about these two examples. First, the evidence in favor of EINI is dubious. Furner says that '[o]ne piece of evidence that Kant's argument here appeals to the will of a hypothetical creator is the phrase 'all capacities...given to him'' (2017b: 320). Furner argues that this should be interpreted as referring to a hypothetical creator because a finite rational being cannot give capacities to itself. Furner offers as a second piece of evidence the phrase 'as a rational being he necessarily wills': Furner says that this should be interpreted as willing in conformity with FLN, which, on the EINI, in turn means willing as a hypothetical creator. But this evidence breaks down almost immediately: the first phrase is referring to an actual creator, not a hypothetical one, and Furner's interpretation of the second phrase is question-begging as evidence.<sup>lii</sup>

Second, it is notable that on Furner's reading both of Kant's examples fail. There would be no contradiction in an EINI hypothetical creator willing the UTC of the natural talents example ('nobody will develop any talents when they need not do so to live a life of comfort') or the UTC of the beneficence example ('nobody will help anybody who needs it when they do not need help themselves to live a life of comfort'). That is because a world in which even a single rational being could live a life of comfort without developing his/her talents would be extravagant, and a world in which any rational being genuinely needs help would be imperfect. Thus, these UTCs *could* be willed by an EINI hypothetical creator because no world willed by such a

creator is extravagant or imperfect and, thus, no possible contradiction can arise: the UTCs are vacuously satisfied in any such world.

Third and finally, the PCI versions of these cases, by way of contrast, do not face these problems. In the natural talents example, the PCI will focus on the fact that the agent's cultivated talents will be serviceable for carrying out his/her will in a way that would be impossible should it be a law of nature that s/he never develop these talents; and in the beneficence example the PCI will focus on the fact that, as a finite rational agent, any human shall will for the help and aid of others at some point in his/her life in a way that would be thwarted should the UTC of his/her maxim of indifference hold as a universal law of nature.<sup>liii</sup>

I would like to conclude this section with some brief remarks about Kleingeld's treatment of the beneficence example. As Kleingeld reconstructs this example, the agent is considering a maxim like, 'I shall never help others in need of assistance, but I shall want others to help me when I need their help to achieve my own ends' (2017: 107). Kleingeld argues that the agent could not will this maxim simultaneously as a universal law and, thus, the maxim fails her VSCA. Furthermore, she argues that this evinces an advantage that the VSCA has over the PCI: whereas the PCI appeals to agents' essential ends to generate a CW, the VSCA 'does not require any such auxiliary assumptions regarding substantive values, essential purposes of the will, or the background conditions of human agency' (2017: 110).

However, the problem with this is that the maxim Kleingeld extrapolates from Kant's text is different from the one conventionally extrapolated. On the conventional reconstruction, the agent's maxim is 'never to help anybody,' and the PCI generates a CW by noting that finite human agents will for the help and aid of others. Kleingeld builds this willing for the help and aid

of others into the original maxim. But this dissolves the advantage she alleges for her account: on the one side, the PCI would not have any difficulty handling Kleingeld's maxim without appealing to essential ends, and on the other, the VSCA does not seem to be able to handle the maxim 'never to help anybody' unless its embargo against such ends is dropped.<sup>liv</sup>

## **Conclusion**

In this paper I sought to defend the traditional interpretations of Kant's FLN from a spate of attempts to refine or reject them, concentrating especially on Furner's CTCLI and EINI. Obviously I have not been able to defend these interpretations (much less FLN itself) against all (or even most) of the textual and philosophical objections leveled against them in the recent literature.<sup>lv</sup> But I do hope that I have shown that these traditional interpretations are more powerful than many recent commentators seem to have supposed and, thus, shown that they are very much still worth grappling with.<sup>lvi</sup>

## Bibliography

'self-murder, n.' OED Online, Oxford University Press, March 2018, [www.oed.com/view/Entry/175356](http://www.oed.com/view/Entry/175356) (accessed 28 May 2018).

'suicide, n.1.' OED Online, Oxford University Press, March 2018, [www.oed.com/view/Entry/193691](http://www.oed.com/view/Entry/193691) (accessed 28 May 2018).

Rivera-Castro, F. 2014. 'Kant's Formula of Universal Law of Nature Reconsidered', *Journal of Moral Philosophy* (11): 185-208. <https://doi.org/10.1163/17455243-4681018>

Furner, J. 2017a. 'Kant's Contradiction in Conception Test', *Theoria* (152): 1-23. <https://doi.org/10.3167/th.2017.6415201>

Furner, J. 2017b. 'Kant's Contradiction in the Will Test', *The Philosophical Forum* (48): 307-323. <https://doi.org/10.1111/phil.12156>

Geiger, I. 2010. 'What is the use of the Universal Law Formula of the Categorical Imperative?' *British Journal for the History of Philosophy* (18): 271-295. <https://doi.org/10.1080/09608781003643568>

Graham, P. 2011. ‘Ought’ and Ability’, *The Philosophical Review* (120): 337-382. <https://doi.org/10.1215/00318108-1263674>

Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge: Harvard University Press.

Kleingeld, P. 2017. ‘Contradiction and Kant’s Formula of Universal Law’, *Kant-Studien* (108): 89-115. <https://doi.org/10.1515/kant-2017-0006>

Kleingeld, P. 2018. ‘A Contradiction of the Right Kind’, *The Philosophical Quarterly* (0): 1-18. <https://doi.org/10.1093/pq/pqy038>

O’Neill (Nell), O. 2013 (1975). *Acting on Principle*. Cambridge: Cambridge University Press (Columbia University Press).

O’Neill, O. 1989. *Constructions of Reason*. Cambridge: Cambridge University Press.

Kahn, S. 2014. ‘The Interconnection between Willing and Believing for Kant’s and Kantian Ethics’, *International Philosophical Quarterly* (54): 143-157. <https://doi.org/10.5840/ipq20145208>

Kant, I. 2011. *Groundwork for a metaphysics of morals*. Ed. and trans. M. Gregor and J. Timmermann. Cambridge: Cambridge University Press.

Korsgaard, C. 1985. 'Kant's Formula of Universal Law', *Pacific Philosophical Quarterly* (66): 24-47. <https://doi.org/10.1111/j.1468-0114.1985.tb00240.x>

Korsgaard, C. 1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.

Nyholm, S. 2015. 'Kant's Universal Law Formula Revisited', *Metaphilosophy* (46): 281-299. <https://doi.org/10.1111/meta.12125>

Wood, A. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.



%20murder%3B%2Cc0%3B%3BSelf%20-%20murder%3B%2Cc0%3B%3BSelf%20-%20Mur-  
der%3B%2Cc0>

<sup>i</sup> GMS, AA 04: 392.3-6. All citations are to the standard Academy pagination of Kant's works. All translations are my own.

<sup>ii</sup> GMS, AA 04: 412.26-414.11.

<sup>iii</sup> GMS, AA 04: 416.7-14.

<sup>iv</sup> GMS, AA 04: 421.18-20.

<sup>v</sup> GMS, AA 04: 424.3-10.

<sup>vi</sup> GMS, AA 04: 421n and GMS, AA 04: 424.10-14.

<sup>vii</sup> *Acting on Principle* was published first in 1975 by Columbia University Press under the name Onora Nell. It was then picked up by Cambridge University Press for its second edition, which came out under Onora O'Neill in 2013.

<sup>viii</sup> 'Kant's Formula of Universal Law' was published first in 1985 as an article in *Pacific Philosophical Quarterly*. Korsgaard then included it in her first anthology, *Creating the Kingdom of Ends*, which came out in 1996.

<sup>ix</sup> 'If universalizing a maxim makes the action proposed inconceivable, then, we can get a logical contradiction' (1996: 82).

<sup>x</sup> This is supposed to be the maxim of the second example Kant uses to illustrate FLN (GMS, AA 04: 422.15-36).

<sup>xi</sup> Rivera-Castro endorses a revised version of the LCI according to which '[t]he contradiction [in the CC test] lies between an assumption in the maxim and its impossibility when the maxim becomes a law' (2014: 203). Her idea seems to be that any given maxim will presuppose various premises, and some of these premises might be false in a world in which the maxim's UTC holds. Thus, on Rivera-Castro's reading, the relevant contradiction is propositional. But I think two points are worth bearing in mind. First, Kant does not talk about the propositions presupposed by the maxims he considers. Second, to read Kant as requiring the CC to manifest propositionally seems anachronistic.

<sup>xii</sup> 'According to the Practical Contradiction Interpretation...the agent would be unable to act on the maxim in a world in which it were universalized so as to achieve his own purpose' (1996: 92).

<sup>xiii</sup> According to Rivera-Castro, 'the practical interpretation understands the question whether I can will a maxim as a universal law as the question whether I can act on the maxim while everyone else adopts it' (2014: 186). But this is incorrect. On the PCI, the question is not whether the act in the maxim is possible when the maxim is a law of nature. Rather, the question is whether the act still will lead to the intended purpose in the maxim when the maxim is a law of nature. The test Rivera-Castro proposes here is more like the LCI. Compare notes 9 and 12 of this paper.

<sup>xiv</sup> 'Willing universal violation creates an inconsistency by making the action-type that it universalizes a non-existence one, and *ipso facto*, ineffectual' (1996: 85). However, it should be noted that in this passage Korsgaard is speaking specifically of conventional actions.

<sup>xv</sup> 'According to the Teleological Contradiction interpretation...we are to consider whether we could will the universalized maxim as a possible law in a teleologically organized system of nature' (1996: 87).

<sup>xvi</sup> GMS, AA 04: 421.7-8. This is from section II of the *Groundwork for a metaphysics of morals*. However, FUL is first introduced in section I (GMS, AA 04: 402.7-9).

<sup>xvii</sup> But see (Kahn 2014: section 1.1).

<sup>xviii</sup> Perhaps Kleingeld would respond that the converse of (2) is false. This would rescue the simultaneity condition. But this response is *prima facie* untenable. And regardless, Kleingeld's claim that the VSCA contradiction is produced by but not equivalent to the LCI and the PCI contradictions entails that the VSCA will not be able to divest itself of any of the presuppositions of these interpretations. I return to this in section 4 below.

<sup>xix</sup> It will be noted that I say nothing about the TCI here. But proponents of the TCI might emphasize the differences between FUL and FLN: although the two are supposed to be equivalent in the sense that they generate the same results for the same maxims, that does not entail that they are applied to these maxims in exactly the same ways.

<sup>xx</sup> In two other formulations in the same chapter O'Neill uses 'at the same time' instead of 'simultaneously' (1990: 83 and 94-95).

That said, it should be noted that Kleingeld mentions willing the maxim as a law of nature whereas O'Neill mentions willing that others adopt the maxim. It thus would be open to Kleingeld to reply that the difference between her account and the traditional interpretations lies in the fact that the latter misunderstand Kant's notion of a law of nature (indeed, both Nyholm and Rivera-Castro attack the traditional interpretations on these grounds; see note 22 below and the paragraph to which it is appended). But such an attack would be distinct from an attack on the basis of the simultaneity condition.

<sup>xxi</sup> A further problem arises when Kleingeld attempts to distinguish the CC test from the CW test:

In the first set of cases [corresponding to the CC test], the volitional self-contradiction is generated by the fact that the maxim cannot be a universal law, because this impossibility makes it contradictory to will the maxim as one's own maxim and simultaneously to will it as a universal law. In the second set of cases [corresponding to the CW test], it is not impossible for the maxim to function as a universal law, but given this maxim's content, willing the maxim contradicts willing what is contained in the idea of a world in which the maxim is a universal law. (2017: 112)

Kleingeld seems to get the text backward here: she has the CW test based on a contradiction in the 'idea' of a world in which the maxim is a universal law (which sounds like Kant's CC test, tracking maxims that cannot be thought as laws of nature); and she has the CC test based on willing a maxim as a universal law (which sounds like Kant's CW test, tracking maxims that cannot be willed as such).

<sup>xxii</sup> On the traditional interpretations, a law of nature is taken to be a causal law. Nyholm bases his interpretation on a text in which Kant distinguishes between a formal and a material use of the term 'nature'. But 'law of nature' is a compound term in the corresponding German (*Naturgesetz*), a compound term that Kant uses regularly in both his theoretical and his practical philosophy. So rather than looking at how Kant uses its component parts, I think a better strategy would be to look at how Kant uses the compound term. Unfortunately, I think this latter strategy tells in favor of the traditional interpretations rather than Nyholm's.

Rivera-Castro also attacks the traditional interpretations on the basis of how they understand the notion of a law of nature. But Rivera-Castro does so from a different angle. Rivera-Castro argues that the traditional interpretations are based on a 'universality as universal adoption' understanding of a law of nature whereas she advocates 'universality as necessity' (2014: 206).

Universality as universal adoption is just that: a maxim is a law of nature on this view if everyone adopts it. Universality as necessity, by way of contrast, involves thinking of a maxim 'as containing a ground...on which the will necessarily determines itself to procure the production of effects' (2014: 199).

The problem with Rivera-Castro's attack, I think, is that she has caricatured the traditional interpretations. As pointed out at the beginning of this note, according to these interpretations elevating the UTC of a maxim to the status of a law of nature involves taking there to be a causal (and, thus, necessary) connection between the action in the maxim and its ends (Wood 1999: 79-81).

<sup>xxiii</sup> However, I think that there are also problems with the first two steps of Nyholm's interpretation. For example, in the second step Nyholm neglects a third motive source that plays an important role on Kant's account: individual inclinations that might conflict with happiness (see, e.g., GMS, AA 04: 399.3-26). See note 22 above for a problem with the first step.

<sup>xxiv</sup> KpV, AA 05: 113.15-23.

<sup>xxv</sup> GMS, AA 04: 404.1-7.

xxvi Given that (as noted in section 1 of this paper) the test of the PCI is stronger than the LCI (in the sense that any maxim that fails the latter *a fortiori* will fail the former), one might conjecture that including the LCI is superfluous and, thus, that Furner should have called his interpretation the CTPCI rather than the CTLCI. However, Furner protests that 'it is false that...[by virtue of this addition, the CTLCI] can no longer be properly distinguished from another interpretation of the CC test, PCI' (2017a: 10). In support of this, he says that on his CTLCI there is a logical contradiction in the conception of a world in which an agent achieves his/her purpose and in which the agent's maxim is a UTC. By way of contrast, he asserts that on Korsgaard's PCI there is 'a specifically practical sense of contradiction' that manifests in an incoherence in an agent's willing a maxim and its UTC (2017a: 11; quoting from Korsgaard).

But I think Furner has misunderstood the difference between the LCI and the PCI. As Korsgaard points out in explaining the PCI, 'it is what happens to the purpose that is the key to the contradiction' (1996: 92). That is, what is key to the PCI is that the agent's purpose (rather than his/her action) is impossible in the world in which the agent's maxim is a UTC; that is the specifically practical sense of contradiction. Whether this idea is explicated in terms of a logical contradiction in the conception of such a world or whether this idea is explicated in terms of an incoherence in the agent's willing his/her maxim and its UTC is inessential. Compare: the key to the LCI is that the agent's action is impossible in the world in which the agent's maxim is a UTC. This, too, could be explicated in terms of a logical contradiction in the conception of such a world or in terms of an incoherence in the agent's willing his/her maxim and its UTC.

So I think that Furner should have called his interpretation the CTPCI rather than the CTLCI. Additional support for this might be garnered from reflection on what a specifically teleological sense of contradiction would be for the TCI.

xxvii Furner goes on to reiterate and defend this feature of the CTLCI in the next section of his article (2017a: 9-11).

xxviii In saying this, Furner is disagreeing with McNair. McNair argues that Furner's stronger requirement ('all' rather than 'some') is too lax. To illustrate why McNair says this, Furner appeals to the following maxim: 'when my view of a lecture is impaired, I will stand up, to get a better view' (2017a: 8). Furner points out that there is a hypothetical system of nature (namely, one in which 'the long-legged always sit behind the short-legged') in which the UTC of this maxim would not create a contradiction even though (i) it would create one in a world like ours and even though (ii) this maxim is intuitively impermissible. From this it may be seen that Furner's stronger requirement runs the risk of resulting in false positives.

In response, Furner objects that appealing to a contradiction in some (rather than all) hypothetical worlds in which the maxim is a UTC fails 'to capture the spirit of the *Groundwork*' (2017a: 9). Furner's argument for this is that different agents might test the maxim as a UTC in different hypothetical worlds and, thus, get different results. This would mean that the reason to refrain from adopting this maxim would not be valid for every rational being, and it would make the hypothetical system of nature 'overly empirical' (2017a: 9).

One thing that is striking about this exchange is that McNair's objection is philosophical whereas Furner's response is textual. The two do not seem to be connecting.

xxix Korsgaard's approach most likely would be an uphill battle for Furner given the suspicion with which people these days seem to regard teleological approaches to nature.

xxx Indeed, when faced with philosophical objections Furner seems simply to double down on the basis of exegetical considerations, perhaps unaware of the problems with which he is saddling Kant (see, for example, note 28 above).

xxxi The latinized 'suicide' does appear in Kant's work (MS, AA 06: 421.16). But it does not appear in the *Groundwork for a metaphysics of morals*, and when Kant uses the term 'Selbstmorde' at GMS, AA 04: 429.16, Timmermann favors translating it as 'self-murder' rather than 'suicide.' Timmermann's argument is: it is 'the natural and literal translation' and '[t]he moralized term 'self-murder' thus seems preferable' (Kant 2011: 169).

For my part, considering the etymology of the term 'suicide' and considering that these terms are synonymous according to the OED entries for them, 'suicide' seems like a perfectly natural and literal translation. And given the history of criminalizing and moralizing suicide (by that name in English speaking countries), a history which has sparked recent movements among proponents of physician assisted suicide to advocate a name-change to 'physician assisted death,' I think that the term 'suicide' is in fact moralized, and it is probably more moralized than the term 'self-murder,' which, to judge from the OED's indication regarding frequency-of-use and from Google's Ngram Viewer (see appendix), is not idiomatic (and, thus, unlikely to evoke a strong affect). So I think that Timmermann's rationale is questionable.

xxxii GMS, AA 04: 422.4-7.

xxxiii GMS, AA 04: 422.8-11.

xxxiv 'They must both end their life...and preserve their capacity for sensation...It is impossible for an agent to act on both of these moral obligations, however. In ending one's life, one destroys all one's capacities, including one's capacity for sensation' (2017a: 14).

xxxv This defangs Furner's point that the PCI and traditional versions of the LCI have a hard time generating a contradiction for self-regarding natural actions like that in the suicide case (2017a: 12). A similar problem underlies Rivera-Castro's attempt to handle this case:

Despite the logical and the practical interpretations, the problem here is not that the action or the purpose in the maxim (ending one's life) would become inconceivable or unattainable. The problem, instead, is that the suicide cannot conceive of his maxim as a law of nature because it contradicts something else he has to consider, namely, that the natural function of self-love is to promote the furtherance of life. (2014: 200)

The reason Rivera-Castro's argument does not work is that if the natural function of self-love were to promote the furtherance of life, then the suicide cannot act on his/her maxim. To put this another way, Rivera-Castro faces a dilemma. Either this premise about the function of self-love is granted to the traditional interpretations, in which case the suicide maxim fails them, or this premise is not granted to Rivera-Castro (because it is too strong), in which case a CC is no longer forthcoming from the suicide maxim on her account.

xxxvi The problem is compounded by the fact that the TCI seems to handle the suicide case with ease. As Furner points out, according to Paton (the most prominent defender of the TCI) the suicide maxim fails the CC test 'because it is contradictory to suppose both that the instinct of self-love always serves the purpose of self-preservation and that the instinct sometimes leads agents to decide to stop living' (2017a: 15), which seems to be parallel with what Kant says in the text.

Furner objects to the TCI reading of the suicide case on the grounds that '[t]he premise about the purpose of self-love is a premise about us' (2017a: 16). That is, self-love might have a different purpose in some other hypothetical world, a purpose that is consistent with the suicide maxim. But I think there are three things that are of note here.

First, Kant often identifies mental capacities and instincts functionally. So self-love, it might be argued, just is the instinct whose function is the promotion of life. If this is correct, then there is no hypothetical world in which the suicide maxim does not fail the TCI of the CC test.

Second, Furner's objection sets him at odds with himself. Even if self-love is not identified functionally, Furner thinks that the requirement that the CTLCI generate a contradiction in *all* possible worlds (rather than merely some) is nonessential (see note 28 above and the paragraph to which it is appended). But the fact that self-love might have a different purpose in *some* other hypothetical world, if it is a fact, would be irrelevant to someone who rejects this nonessential feature of the CTLCI.

Third and perhaps most importantly, Furner is engaged in a purely exegetical project. With that in mind, Paton might reply simply that Furner's objection is not germane. Indeed, precisely this renders Furner's interpretation of the suicide case tenuous: nowhere in Kant's discussion does he mention the 'necessary means' for acting on the suicide maxim, nor does he say that these means include preserving the capacity for sensation. Kant baldly introduces the claim that the purpose of self-love is the promotion of life; this is not something inferred about the means for committing suicide from self-love, and it is not the promotion of self-love itself but of the creature who has this instinct.

xxxvii GMS, AA 04: 403.16-17, quoted on Furner 2017a: 18.

xxxviii GMS, AA 04: 422.34, quoted on Furner 2017a: 18.

xxxix It also is relevant that the quotation Furner uses to back up his claim about timing is from Kant's discussion of FUL (not FLN) in part I (not part II) of the *Groundwork for a metaphysics of morals*. This is relevant because some might assert that the test associated with FLN is different from that associated with FUL (see note 19 above).

<sup>xl</sup> Such suspicions might be thought to lead us back to the CTLCI in light of what I said at the beginning of this paragraph in criticizing (A). But given the textual problems already documented in my discussion of 1 and 2, I do not think this is sustainable. Further problems might arise from the at least *prima facie* disunity between CTCLI and Furner's interpretation of the CW test (discussed in the next section of this paper).

<sup>xli</sup> However, in parting I point out that Furner neglects to discuss the piece of textual evidence that Korsgaard thinks tells most strongly in favor of the PCI, Kant's treatment of the hypothetical imperative (Korsgaard 1996: 93-94).

<sup>xlii</sup> One might distinguish between the CI and what it enjoins us to do, the many ci's that follow from the CI: I ought to act as if my maxims could be willed as universal laws of nature, whence it follows that I should not run off from a restaurant without paying some determinate amount of money (the balance of my bill). The only way to bridge the gap is by appeal to external, empirical premises, and that is not a problem for Kant, for the bindingness or *priority* of his moral law or for the PCI.

<sup>xliii</sup> I think it bears emphasis again (as emphasized in the opening paragraphs of both section 3 and section 4 of this paper) that Furner's project is strictly textual. If Furner's project were philosophical then I think he would be open to a different kind of objection here.

<sup>xliv</sup> MS, AA 06: 241.20-21. This comes out especially in the first and seventh convolutes of the *Opus postumum*: Kant maintains there that we should think of God as all-obliging but never obliged, and he uses this to argue for the uniqueness of God (e.g., OP, AA 22: 124.22-26).

<sup>xliv</sup> MS, AA 06: 451.4-5.

<sup>xlvi</sup> TP, AA 08: 278.16-18. A more controversial example can be drawn from Kant's doctrine of the highest good in the *Critique of the power of judgment*.

In earlier articulations of this doctrine, Kant says that there is a duty to promote the highest good and, thus, if the highest good is impossible, the moral law is invalid (KpV, AA 05:114.6-9). This, according to Kant, generates moral grounds for belief in God because the highest good is really possible only if God exists.

But in the *Critique of the power of judgment* Kant suggests that if an individual is unable to convince him/herself that God exists, then s/he would not have to give up on the moral law *tout court*. Rather, s/he merely would have to give up on the duty to promote the highest good (KU, AA 05:450.31-451.7). This suggests that, according to Kant, empirical information about individual agents (like whether a given individual believes in God) can factor into the determination of our duties.

<sup>xlvii</sup> It is perhaps also worth pointing out that the EINI is also going to require external premises: how can we determine whether a nature is extravagant or imperfect (as Furner understands these terms) without knowledge of the ends at which the beings in this nature aim?

<sup>xlvi</sup>iii GMS, AA 04: 423.1-2.

<sup>xliv</sup>x GMS, AA 04: 423.13-16.

<sup>l</sup> GMS, AA 04: 423.17-18.

<sup>li</sup> GMS, AA 04: 423.33-35.

<sup>lii</sup> It is also strained: even if the EINI were correct, it would be as an actual rational being that s/he wills that all his/her capacities be developed.

<sup>liii</sup> Furner raises two objections against this. First, he maintains that on the PCI, Kant overstates his case in the natural talents example: a rational agent necessarily would will to develop only *some* rather than *all* of his/her talents (2017b: 321). Second, Furner claims that on the PCI Kant's beneficence example proves too much because situations might arise in which agents urgently want to 'go it on their own,' and reasoning parallel to that used in the PCI version of the case thereby 'would defeat any maxim of helping others in distress' (2017b: 322).

But I do not think either of these objections works. Against the first objection, I note that as a matter of pre-theoretic intuition, it seems to me that Kant did overstate his case: a duty to develop *all* of one's talents would be incoherent. Against the second objection, I concede that parallel reasoning would defeat a maxim of *always* helping others in distress. But for one thing, this probably agrees with intuition: paternalistically helping others even when others do not want help probably *is* impermissible. And for another thing, parallel reasoning would *not* defeat a more modest maxim of helping when so desired.

<sup>liv</sup> A similar problem underlies Rivera-Castro's treatment of maxims that contain natural actions. According to Rivera-Castro, her revised version of the LCI 'can account for the immorality of maxims that contain natural actions, while neither the practical nor the standard logical interpretation can do so because such actions do not become either frustrated or inconceivable upon universalization' (2014: 205). To illustrate her point, Rivera-Castro uses two examples: (1) a maxim of convenience killing and (2) a maxim of convenience torturing. Rivera-Castro contends that an agent with a maxim of convenience killing assumes that 'she will stay alive while she eliminates other people in order to obtain whatever purpose she is pursuing,' and an agent with a maxim of convenience torturing makes a similar assumption about remaining free from torture (2014: 205). Because the UTCs of these maxims would render these assumptions false, Rivera-Castro asserts that the maxims fail her revised version of the LCI (see note 11 above).

However, Rivera-Castro fails to notice that if the maxims are constructed with purposes that require the agent to stay alive or be free from torture, the UTCs of these maxims *would* generate a CC from the PCI (or to be more precise: they would generate a CC from the PCI if but only if they would generate a CC from Rivera-Castro's revised version of the LCI).

<sup>lv</sup> For example, I have not said anything about what Kleingeld calls the Latitude Problem. The Latitude Problem is that the maxim of convenience killing 'seems to generate the *wrong kind* of contradiction for showing convenience killing to be absolutely impermissible' (2018: 2). That is, this maxim generates a CW, but Kant says that maxims that fail the CW correspond to imperfect duties and that imperfect duties allow a degree of latitude in their fulfillment.

Kleingeld argues that the Latitude Problem rests on a subtle equivocation in the following argument: (1) non-beneficence is sometimes permissible; (2) convenience killing is sorted with non-beneficence by FLN; therefore (3) convenience killing is sometimes permissible. According to Kleingeld, (1) is true but (2) is false if non-beneficence is understood as *not acting on one's principle of beneficence*, whereas (1) is false but (2) is true if non-beneficence is understood as *principled non-beneficence*. In this way Kleingeld attempts to show that convenience killing can be absolutely impermissible notwithstanding the fact that it generates a CW rather than a CC.

I would like to make three quick comments about this. First, I think there might be room for proponents of the traditional interpretations to argue that the maxim of convenience killing generates a CC (see note 54 above).

Second, in formulating the Latitude Problem as she does, Kleingeld overlooks a crucial aspect of the trouble that arises for FLN if convenience killing generates a CW rather than a CC: this would seem to entail that convenience killing is a less serious fault than telling a lying promise, a result that is *prima facie* counterintuitive (see Herman, 1993, 118). Relatedly, Kant's so-called Formula of Humanity formulation of the CI seems to make convenience killing a more serious offense, thus creating problems for the purported equivalence of these two formulations.

Third and finally, I think Kleingeld's attempt to smooth away Kant's claim that perfect duties allow no exceptions in favor of inclination is questionable (2018: 10n12). However, a detailed discussion of this is beyond the scope of this paper.

<sup>lvi</sup> I would like to express my deepest gratitude to Allen Wood: even where I disagree with him about how to interpret FLN, I only have been able to understand my own views (to the limited extent that I do) thanks to his incisive and generous conversation and correspondence over the years.